



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Breast cancer risk reduction and membrane-bound catechol O-methyltransferase genetic polymorphisms

Citation for published version:

Ji, Y, Olson, J, Zhang, J, Hildebrandt, M, Wang, L, Ingle, J, Fredericksen, Z, Sellers, T, Miller, W, Dixon, JM, Brauch, H, Eichelbaum, M, Justenhoven, C, Hamann, U, Ko, Y, Brüning, T, Chang-Claude, J, Wang-Gohrke, S, Schaid, D & Weinshilboum, R 2008, 'Breast cancer risk reduction and membrane-bound catechol O-methyltransferase genetic polymorphisms', *Cancer Research*, vol. 68, no. 14, pp. 5997-6005. <https://doi.org/10.1158/0008-5472.CAN-08-0043>

Digital Object Identifier (DOI):

[10.1158/0008-5472.CAN-08-0043](https://doi.org/10.1158/0008-5472.CAN-08-0043)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Cancer Research

Publisher Rights Statement:

Published in final edited form as:
Cancer Res. 2008 July 15; 68(14): 5997–6005.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Published in final edited form as:

Cancer Res. 2008 July 15; 68(14): 5997–6005.

Breast Cancer Risk Reduction and Membrane-Bound Catechol O-Methyltransferase Genetic Polymorphisms

Yuan Ji^{1,*}, Janet Olson^{2,*}, Jianping Zhang^{1,*}, Michelle Hildebrandt¹, Liewei Wang¹, James Ingle³, Zachary Fredericksen², Thomas Sellers⁴, William R. Miller⁵, J. Michael Dixon⁵, Hiltrud Brauch⁶, Michel Eichelbaum⁶, Christina Justenhoven⁶, Ute Hamann⁷, Yon Ko⁸, Thomas Brüning⁹, Jenny Chang-Claude⁷, Shan Wang-Gohrke¹⁰, Daniel Schaid², and Richard Weinshilboum¹

¹Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic College of Medicine, Mayo Clinic, Rochester, MN ²Department of Health Sciences Research, Mayo Clinic College of Medicine, Mayo Clinic, Rochester, MN ³Department of Oncology, Mayo Clinic College of Medicine, Mayo Clinic, Rochester, MN ⁴H. Lee Moffitt Cancer Center, Tampa, FL ⁵Western General Hospital, Edinburgh, Scotland ⁶Dr. Margarete Fischer-Bosch-Institut für Klinische Pharmakologie, Stuttgart, and Tübingen Universität, Tübingen, Germany ⁷German Cancer Research Center, Heidelberg, Germany ⁸Johanniter Hospital Bonn, Bonn, Germany ⁹Berufsgenossenschaftliches Forschungsinstitut für Arbeitsmedizin, Ruhr Universität Bochum, Germany ¹⁰Women's Hospital of the University of Ulm, Ulm, Germany

Abstract

Catechol O-methyltransferase (COMT)-catalyzed methylation of catecholestrogens has been proposed to play a protective role in estrogen-induced genotoxic carcinogenesis. We have taken a comprehensive approach to test the hypothesis that genetic variation in *COMT* might influence breast cancer risk. Fifteen *COMT* SNPs selected on the basis of in-depth resequencing of the *COMT* gene were genotyped in 1482 DNA samples from a Mayo Clinic breast cancer case-control study. Two common SNPs in the distal promoter for membrane-bound (MB) COMT, rs2020917 and rs737865, were associated with breast cancer risk reduction in premenopausal women in the Mayo Clinic study, with allele-specific odds ratios of 0.70 (95% CI = 0.52–0.95) and 0.68 (95% CI = 0.51–0.92), respectively. These two SNPs were then subjected to functional genomic analysis and were genotyped in an additional 3683 DNA samples from two independent case-control studies (GENICA and GESBC). Functional genomic experiments showed that these SNPs could up-regulate transcription and that they altered DNA-protein binding patterns. Furthermore, substrate kinetic and exon array analyses suggested a role for MB-COMT in catecholestrogen inactivation. The GENICA results were similar to the Mayo case-control observations, with ORs of 0.85 (95% CI = 0.72–1.00) and 0.85 (95% CI = 0.72–1.01) for the two SNPs. No significant effect was observed in the GESBC study. These studies demonstrated that two SNPs in the *COMT* distal promoter were associated with breast cancer risk reduction in 2 of 3 case-control studies, compatible with the results of functional genomic experiments, suggesting a role for MB-COMT in breast cancer risk.

Address correspondence to: Richard M. Weinshilboum, M.D., Division of Clinical Pharmacology, Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN 55905 USA, Telephone: (507) 284-2246; Fax: (507) 284-4455; E-mail: Weinshilboum.Richard@mayo.edu.

*The first three authors contributed equally to the work described in this manuscript.

Keywords

Catechol O-methyltransferase; COMT; MB-COMT; S-COMT; breast cancer risk; genetic polymorphism; SNPs; functional genomics

Introduction

Breast cancer is the most common cancer and the second leading cause of cancer death in women (1). Life-time estrogen exposure is a major risk factor for breast cancer (2,3). Estrogens can potentially induce carcinogenesis through both receptor-mediated and non-receptor mediated pathways (4–6) (Fig. 1). This latter pathway is thought to involve metabolic activation, during which estrogens are metabolized by cytochrome P450 enzymes to form catecholestrogens which can then be converted to estrogen quinones capable of forming stable or depurinating DNA adducts, resulting in genotoxicity (5,7). Since 4-OH catecholestrogens are also potent ER ligands (8,9), catecholestrogens could potentially influence both pathways for estrogen carcinogenesis. Catechol O-methyltransferase (EC2.1.1.6; COMT) catalyzes the methylation of catechols, including catecholestrogens (10,11). In addition, 2-methoxy catecholestrogens also have strong anti-carcinogenic effects (12). Therefore, decreased COMT activity might be a risk factor for breast cancer while elevated activity would, in theory, be protective.

COMT encodes two isoforms, a soluble cytoplasmic (S-COMT) and a membrane-bound isoform (MB-COMT) (14,15) (Fig. 2A). A “proximal promoter” in intron 2 drives S-COMT transcription, while a “distal promoter”, 20 kb upstream at the 5′-end of the gene, does the same for MB-COMT (14,15). The two COMT isoforms differ only by an additional 50 hydrophobic amino acids at the N-terminus of MB-COMT (14,15). MB-COMT is believed to be the dominant isoform in the brain, while S-COMT predominates in peripheral tissues (14–16).

We first described the autosomal co-dominant inheritance of S-COMT activity 30 years ago (17,18), and its genetic basis was subsequently shown to result, in part, from a nonsynonymous G to A polymorphism that changes a Val to Met at codon 108 for S-COMT and codon 158 for MB-COMT (19). The Met108/158 allele is associated with decreased enzyme activity and decreased thermal stability, both *in vitro* and *in vivo* (17,18,20–22). During the past decade, over 40 epidemiologic studies have been performed to test the possible association of *COMT* sequence variation with breast cancer risk. Nearly all of those studies have focused on only the Val108/158Met polymorphism, but the results have been inconsistent (23–25).

In the present study, we took a systematic approach to test the hypothesis that genetic polymorphisms in *COMT* might influence breast cancer risk. We began by resequencing the gene to determine the spectrum of common variation in *COMT* (Fig. 2A). Fifteen *COMT* polymorphisms were then chosen to genotype DNA samples from a large Mayo Clinic breast cancer case-control study – followed by functional studies of polymorphisms that appeared to be associated with risk. We then performed two separate replication studies to test the observed association. This comprehensive approach resulted in the identification of two polymorphisms in the distal promoter of *COMT* that were associated with decreased breast cancer risk in two of the three studies – an association strongly supported by the functional genomic results – suggesting that MB-COMT might play a role in catecholestrogen metabolism in breast cancer. Therefore, our study strongly suggests that these MB-COMT SNPs may represent another of an increasing number of common polymorphisms that influence breast cancer risk (26).

Materials and Methods

SNP Selection for Genotyping

COMT was resequenced using DNA from 60 Caucasian-American subjects for all exons, splice-junctions, intron 2 (22) and 1 kb on either side of exon 1¹. The *COMT* consensus sequence used in our studies was that of contig NT_011519.10. Numbering of nucleotides within the coding region began at the “A” in the “ATG” for MB-COMT, with nucleotides 3’ to that position assigned positive numbers. Nucleotides located within the two upstream non-coding exons, and the 5’-FR were assigned negative numbers, with the final nucleotide of exon 2 designed “-1” (i.e., the nucleotide 5’ of the “ATG” start codon). Nucleotides located within introns were numbered on the basis of distance from the nearest splice site, using positive and negative numbers for distance to 5’- and 3’-splice sites, respectively. The numbering scheme for amino acids was based on their distance from the MB-COMT translation start codon. SNPs included in the genotyping studies were selected either by use of haplotype-tagging (27) or the LD-select method (28). Fifteen SNPs were selected in this fashion or on the basis of previous reports of their association with clinical phenotypes (29,30).

Study Populations

Mayo Clinic breast cancer case-control study—The Mayo Clinic study is an unselected, clinic-based series of breast cancer cases and healthy controls recruited in the Mayo Clinic Department of Oncology (31). Cases had a diagnosis of histologically confirmed primary breast cancer within 6 months of enrollment and no prior history of cancer (except for non-melanoma skin cancer). Controls were frequency matched to cases on county of residence and five-year age group and were selected from women visiting the Mayo Clinic for a general medical examination. “Postmenopausal” was defined as having no menstrual period for 12 months or having had the uterus and/or ovaries removed. Family history of cancer was reported for all first and second degree relatives. Controls were ineligible if they had a previous diagnosis of cancer (except non-melanoma skin cancer). As of June 2005 when the present study began, 750 cases and 732 controls, entirely Caucasian, had been recruited from the states of Iowa, Minnesota or Wisconsin. Participation rates were 70% among cases and 72% among controls. All eligible women were asked to provide written informed consent, risk factor information via a written questionnaire, and a sample of blood. The study was reviewed and approved by the Mayo Clinic Institutional Review Board.

The Interdisciplinary Study Group on Gene Environment Interactions and Breast Cancer in Germany Study Population (GENICA Study)—The GENICA study is a study of breast cancer cases and controls from the greater Bonn area in Germany (32). Specially, between August 2000 and September 2004, 1143 breast cancer cases and 1155 population-based controls were recruited. Cases had a first-time diagnosis of primary breast cancer that was histologically confirmed within 6 months of enrollment. Controls were population-based and matched to cases in 5-year age groups. All participants were Caucasian and were <80 years of age. Risk factor information was collected via in-person interviews using the core questionnaire of a German population survey that had been extended by the inclusion of questions on factors related to breast cancer. The response rates were 86% for cases and 69% for controls. A participant was classified as premenopausal if she reported menstrual periods in the year of interview. All participants provided a blood sample, and DNA was available for 1015 cases and 1021 controls. The GENICA study was reviewed and approved by the Ethics Committee of the University of Bonn; and all study participants provided written informed consent.

¹J. Zhang, et al. Catechol O-Methyltransferase Pharmacogenomics: Human Liver Genotype-Phenotype Correlation and Proximal Promoter Studies. Submitted to “Pharmacogenetics and Genomics”.

The Genetic Epidemiology Study of Breast Cancer by Age 50 (GESBC)—The GESBC is a population-based case-control study carried out in two geographical areas of Germany (Freiburg and the Rhein-Neckar-Odenwald regions) that included patients newly diagnosed with breast cancer between 1992 and 1995 (33). Subjects were German speaking (97% had two, and 3% one German parent), under 51 years of age, and they could have no previous history of breast cancer. Cases had a confirmed diagnosis of either *in situ* or invasive breast cancer and were identified by monitoring approximately 40 hospitals. Participation was 70% among the 1005 eligible and living cases. An attempt was made to recruit two population controls per case, matched by age and study region. Subjects were not eligible to be controls if they could not speak German, had moved out of the study region, or had a previous history of primary breast cancer. Participation was 61% among the 2257 eligible population controls. Subjects were defined as postmenopausal if they reported a natural menopause 6 months before the reference date (date of diagnosis or date of questionnaire for control) or a bilateral oophorectomy. The menopausal status of women who reported hysterectomy alone was classified as unknown. Subjects were asked to complete a self-administered questionnaire on demographic and breast cancer risk factors, and to provide a blood sample. The GESBC study was reviewed and approved by the Ethics Committee of the University of Heidelberg, and subjects who participated provided written informed consent.

Genotyping Methods

Fifteen *COMT* polymorphisms were used to genotype samples from the Mayo Clinic study. Most genotyping was performed with the SNPstream platform (Beckman Coulter, Fullerton, CA). ABI fragment analysis assay was used to genotype the insertion/deletion polymorphism at *COMT* 3'-UTR(820), and DNA sequencing was used to genotype five SNPs: 5'-FR(-1420), 5'-FR(-628), 5'-FR(-485), I2(832) and I2(1140). During the two replication studies, only 3 SNPs, 5'-FR(-628), I1(701) and the exon 4(472) Val108/158Met polymorphism, were genotyped. The GENICA samples was genotyped using Sequenom MALDI-TOF MS (34). TaqMan technology was used to genotype the GESBC samples (35).

Statistical Analysis

Demographic information for cases and controls was summarized using means and standard deviations, or counts and percentages, as appropriate. Comparisons were made between case and control groups using t-tests or chi-square tests. Genotypes for each of the polymorphisms were used to estimate allele frequencies separately for cases and controls. Genotypes in the controls were also assessed for departures from Hardy-Weinberg equilibrium.

For the Mayo Clinic study, the association of each SNP genotype with case/control status was performed separately for pre- and postmenopausal strata. Within each stratum, unconditional logistic regression was used to evaluate the log-additive effects of the rare allele for each SNP, while adjusting for age and geographical region of residence. The log-additive effects were summarized by odds ratios (ORs), and 95% confidence intervals (95% CIs).

In addition to the matching variables, the effects of other covariates were evaluated using backward stepwise regression, and the additional significant covariates (p-value < 0.05) were parity/age at first child birth and physical activity. These additional covariates were also used for statistical adjustments in the logistic regression analyses. Although there have been a large number of studies of the effect of *COMT* on breast cancer risk, making the need to control for multiple testing for our study debatable, we nonetheless present p-values adjusted for multiple testing of 30 tests (15 SNPs and 2 menopausal strata). To adjust p-values, we randomly permuted case/control status within each menopausal-status stratum 10,000 times to determine the distribution of the test statistics. The frequency that the largest permuted statistic (out of 30) was greater than each of the observed statistics provided an empirical p-value adjusted for

the 30 tests performed. Supplemental analyses were used to evaluate the association of SNP genotypes with estrogen receptor (ER) status by directly comparing the ER(+) cases versus the ER(−) cases, once again using logistic regression with covariates.

For the GENICA and GESBC studies, backward covariate selection was used to choose the covariates HRT and physical activity for the GENICA study and physical activity and parity/age at first birth for the GESBC study. To perform pooled analyses across all three studies, relevant covariates were included as fixed effects and the study was included as a random effect, fitting the models with the SAS NLMIXED procedure. Analyses were performed in the SAS (SAS Institute, Cary, NC), S-Plus (Insightful, Seattle, WA) and Haploview 3.3 (Daly Lab at the Broad Institute, Cambridge, MA) programs.

Functional Genomic Analyses

Electromobility shift (EMS) assays—EMS assays were performed with MCF-7 cell nuclear extract for oligonucleotides containing the wild type (WT) or variant nucleotides (see Supplemental Table 1 for oligonucleotide sequences). The procedure has been described elsewhere (36). ³²P-labeled annealed oligonucleotides were incubated with 5 µg nuclear extract protein (Active Motif, Carlsbad, CA) and 5X binding buffer (Promega Corporation, Madison, WI) for the binding reactions. Nuclear extract was eliminated from the control reaction. One hundred-fold excess unlabeled specific oligonucleotide was added to the competition reactions. All reactions were incubated at room temperature for 20 minutes and were then subjected to PAGE. The AliBaba2.1² program predicted that the 5'-FR(-628) sequence might bind to AP2α. Therefore, a supershift assay was also performed with the WT C(-628) oligonucleotides using a rabbit polyclonal antibody against human AP2α (Santa Cruz Biotechnology, Inc. Santa Cruz, CA).

Reporter gene studies—The C(-628)T and IIA(701)G, SNPs were located within an area containing the *COMT* distal promoter. To study their possible effect on transcription, reporter gene constructs were created by cloning 1.5 kb of DNA sequence spanning the two SNPs into pGL3-Basic (Promega). Two single variants, as well as a double variant with both SNPs, were created from the WT construct by site-directed mutagenesis. Inserts were sequenced in both directions to ensure that the correct sequence was present. These reporter gene constructs were designated as the WT, (-628)C/I1(701)A, (-628)T/I1(701)A, (-628)C/I1(701)G and (-628)T/I1(701)G – the double variant. 1.5 µg of each reporter gene construct was co-transfected with 20 ng of the pRL-TK vector into breast carcinoma MCF-7 cells, MDA468 cells and hepatic carcinoma HepG2 cells, followed by dual-luciferase assay (Promega). The values for relative luciferase activity were reported as percentages of the WT activity.

Recombinant COMT expression and enzyme assays—Expression constructs encoding the WT MB-COMT and S-COMT proteins were generated by cloning the cDNAs into the eukaryotic expression vector pCR3.1 (Invitrogen, Carlsbad, CA). These expression constructs were transfected into COS-1 cells and cytosol (containing S-COMT) and microsome (containing MB-COMT) preparations were isolated (37). COMT activity was measured with 4-hydroxyestradiol and 2-hydroxyestradiol (Sigma-Aldrich, St. Louis, MO) as methyl acceptor catechol substrates (38). Apparent *Km* values for 4-OH-E2 and 2-OH-E2 were determined with a series of substrate concentrations ranging from 0.05 to 20 µM.

COMT exon array analyses—Total RNA was extracted from flash-frozen breast cancer tissue from 7 individual patients. All RNA samples had an Agilent RNA Integrity Number of greater than 7.8. The RNA was then reverse-transcribed and biotin labeled for hybridization

²<http://www.gene-regulation.com/pub/programs/alibaba2/index.html>

with GeneChip® Human Exon 1.0 ST Arrays (Affymetrix, Santa Clara, CA). Arrays were normalized with full quantile normalization using XRAY software (Biotique Systems, Inc., Reno, NV) and were background corrected using GC-content matched antigenomic probes by the application of median-polish (exon RMA). Log2 transformed expression values were derived for each of the *COMT* “core” probe sets.

Results

COMT Haplotype Structure

COMT was resequenced using DNA samples from 60 Caucasian subjects (22). A total of 33 polymorphisms were observed, including the Val108/158Met polymorphism that has been studied so often (Fig. 2A). One third of the polymorphisms were located within the region flanking exon 1, an area that includes the distal promoter for MB-COMT. There appeared to be at least two clearly defined “haplotype blocks” in DNA samples from CA subjects (Fig. 2B).

Mayo Clinic Breast Cancer Case-Control Association Study

The Mayo Clinic study included 750 cases as well as 732 controls. Characteristics of these subjects, as well as the GENICA and GESBC subjects, are listed in Supplemental Table 2. For the Mayo Clinic study, the controls tended to have significantly more live-births than did the cases, and fewer cases were highly physically active when compared with controls. For the GENICA study, cases also tended to be less active than controls; but not in the GESBC study, which recruited younger cases. Finally, cases had a higher frequency of family history of breast or ovarian cancer than did their controls in all three studies. All SNP genotypes fit HWE proportions among the controls (p-values ranging from 0.14–0.94 across all three groups), indicating high-quality genotyping.

Associations of *COMT* SNPs with case/control status for the Mayo Clinic study are summarized in Table 1. Two SNPs, rs2020917[C(-628)T] and rs737865 [I1A(701)G], appeared to be associated with breast cancer risk reduction in premenopausal women. The analysis was stratified for menopausal status because previous studies of *COMT* and breast cancer risk had demonstrated differences between pre- and postmenopausal women (23,25), and because of a great deal of evidence that the pathophysiology of the disease may differ in these two groups – including recent data from genome-wide association studies (39–41). The estimated allele-specific ORs for the rare alleles of both SNPs were approximately 0.70, and adjustment for covariates had little impact on these estimates. Comparing the log-additive model of allele effects to the general model for codominant effects did not demonstrate any statistically significant departures from the log-additive effects. Additional analyses (not presented) suggested that the effects of these two SNPs did not differ significantly between ER(+) and ER(–) cases within each of the menopausal strata (p-values ranging from 0.28 to 0.68). These two SNPs were in high LD, both in breast cancer cases ($D' = 0.98$ and $R^2 = 0.95$) and controls ($D' = 0.91$ and $R^2 = 0.78$). Although the p-values unadjusted for multiple testing suggested that these SNPs provide a reduced risk of breast cancer among premenopausal women (p-value = 0.022 for 5'-FR(-628) and p-value = 0.011 for I1(701)), after correcting for multiple testing, the associations no longer met the usual criteria for statistical significance (p-value = 0.201 and 0.102, respectively). However, given the number of studies that have reported a positive association of *COMT* with breast cancer, we concluded that these two SNPs were worth pursuing in replicate studies, especially considering the results of the functional genomic experiments described subsequently.

The Val108/158Met polymorphism at exon 4(472) that has been studied in all previous *COMT*-breast cancer studies did not appear to be associated with breast cancer risk in the Mayo

study. Both of the distal promoter SNPs, 5'-FR(-628) and I1(701), were moderately linked to the Val108/158Met polymorphism ($D' = 0.78$, $R^2 = 0.21$ and $D' = 0.91$, $R^2 = 0.31$, respectively) in the samples used for our *COMT* resequencing studies.

Functional Genomic Studies

EMS assays and reporter gene studies—The results of the Mayo Clinic study suggested that the two distal promoter SNPs, 5'-FR(-628) and I1(701), might play a role in breast cancer risk reduction. However, before attempting replication, we performed a series of functional genomic experiments to determine the biological plausibility of these SNPs. First, EMS assays were performed with MCF-7 cell nuclear extract. Fig. 3A shows that nuclear protein(s) bound to sequences at or near nucleotide (-628), but the change from C to T at that position did not alter the binding pattern. On the basis of a supershift assay, the protein bound to the (-628) locus might be an AP2 transcription factor (Fig. 3B). At the I1(701) locus, WT oligonucleotides did not bind protein, but protein binding occurred with the variant oligonucleotide sequence (Fig. 3A).

Next, reporter gene studies were performed. Constructs for single variants with each SNP alone as well as a construct containing both variant nucleotides were created using site-directed mutagenesis. The double variant had the greatest effect on transcription when compared with the WT in all three cell lines studied (all p-values <0.002) (Fig. 3C). This effect was particularly striking in MDA468 cells, with an average 2.3-fold increase (p-value <0.0002).

Substrate kinetic and exon array analyses—Both SNPs that appeared to influence breast cancer risk were located in the area containing the promoter for MB-COMT. Although it is known that MB-COMT catalyzes the methylation of catecholamine neurotransmitters with lower *K_m* values than those for S-COMT (42), there is no similar evidence for catecholestrogens, so we performed catecholesterogen substrate kinetic studies. Both 2-OH and 4-OH catecholestrogens were substrates for O-methylation catalyzed by MB-COMT (Fig. 4A). The apparent *K_m* for S-COMT was 16.1 ± 2.0 (mean \pm SEM) μ M for 4-OH-E2 and 9.0 ± 0.66 μ M for 2-OH-E2. MB-COMT had significant lower *K_m* values for both catecholesterogen substrates, 6.1 ± 0.55 μ M for 4-OH-E2 and 1.0 ± 0.04 μ M for 2-OH-E2 (p < 0.05 compared with *K_m* values for S-COMT). We also performed exon array analyses with RNA from breast cancer tissue to determine whether MB-COMT was expressed in that tissue. Exon array analyses performed with total RNA extracted from 7 breast cancer tissue samples (Fig. 4B) showed that the two non-coding upstream exons, exons present only in MB-COMT mRNA, were expressed in all breast tumor samples.

Results from these functional genomic studies supported a role for MB-COMT and the two distal promoter SNPs identified during the Mayo Clinic study. This evidence of biological plausibility provided a stimulus for the replication studies described subsequently.

Replication Studies

The two replication studies performed in collaboration with GENICA and GESBC involved genotyping only three SNPs (thus reducing multiple testing), the heavily studied Val108/158Met – exon 4(472) polymorphism and the two distal promoter SNPs. The association of these SNPs with case-control status in the GENICA and GESBC samples is illustrated in Table 2. For the GENICA study, the 5'-FR(-628) and I1(701) SNPs were significantly associated with a reduced risk for breast cancer in the postmenopausal group. Although the associations of these SNPs in the premenopausal stratum were not statistically significant by the usual criterion of p-value < 0.05, the sample sizes were much smaller for the premenopausal stratum, and the magnitude of the allelic effects for the GENICA study were consistent between the pre- and postmenopausal strata, as well as with the risk reduction

observed in the Mayo Clinic premenopausal stratum. In contrast, the GESBC study did not show a risk reduction for these SNPs. In fact, the ORs for that study were slightly greater than 1.0.

Finally, in an attempt to evaluate effect sizes pooled across all three studies, we allowed the covariates to differ for the three studies (each study was adjusted for its own set of covariates), but estimated a pooled effect size of the allelic OR for each of the SNPs (see Supplemental Table 3). Two different p-values are listed: the p-value for OR = 1 tests whether the pooled OR differs significantly from 1.0, while the p-value for heterogeneity tests whether the ORs differ across the three studies. The ORs for 5'-FR(-628) and I1(701) differed significantly among the three studies for the premenopausal stratum. Tables 1 and 2 show that these SNPs were associated with reduced risk for both the Mayo Clinic and GENICA studies in at least one of the menopausal cohorts, but displayed no effect in either cohort for the GESBC study. In contrast, none of the ORs differed significantly in the postmenopausal stratum.

Discussion

Known breast cancer susceptibility genes, including *BRCA1* and *BRCA2*, account for only a small portion of the familial risk for breast cancer, resulting in many studies with the goal of identifying low-penetrance risk genes (43). However, most of those efforts were unsuccessful prior to a recent series of large genome-wide association studies (26,40,41). Those studies identified common polymorphisms which might represent novel markers for breast cancer risk. The largest of those studies concluded that many additional common susceptibility alleles remained to be identified (26). Estrogen exposure is an established breast cancer risk factor (2,3). COMT could potentially reduce risk for estrogen-dependent carcinogenesis in both ER-dependent and non-ER-dependent pathways by catalyzing the O-methylation of catecholestrogens (5,7) (Fig. 1). As a result, numerous breast cancer association studies have been performed, almost all focusing on the Val108/158Met polymorphism (23–25), and the results of those studies have been inconsistent.

In the present study, we took a comprehensive approach that began with resequencing of the *COMT* gene (Fig. 2A). Fifteen of the 33 polymorphisms identified in DNA from Caucasian subjects were used to genotype DNA samples from a large Mayo Clinic breast cancer case-control study. Those 15 polymorphisms were scattered across the entire length of the gene within both major haplotype blocks present in Caucasian subjects (Fig. 2B). We observed a reduction in breast cancer risk in premenopausal women associated with two SNPs in the “distal promoter” for MB-COMT, but the heavily studied Val108/158Met polymorphism failed to show a significant association with risk.

Functional genomic studies of these two polymorphisms demonstrated that they altered nuclear protein binding patterns and were associated with the up-regulation of transcription – potentially resulting in increased COMT activity – compatible with reduced risk if the “catecholesterol carcinogenesis hypothesis” is correct. These two SNPs were highly linked, but both were only moderately linked to the Val108/158Met polymorphism, possibly helping to explain why that polymorphism has been inconstantly associated with risk. Fig. 2B shows at least two large haplotype blocks in *COMT*, and 5'-FR(-628) and I1(701) are both located in the upstream block associated with MB-COMT. The location of these two functional polymorphisms in the distal promoter suggested that MB-COMT might play a role in estrogen-induced carcinogenesis. Previous studies were based on the tacit assumption that S-COMT is the major isoform in most peripheral tissues, while MB-isoform is most important in the brain (14,16). However, our substrate kinetic experiments showed that MB-COMT has a higher affinity for catecholestrogens than does S-COMT (Fig. 4A). Furthermore, exon array analysis for 7 breast cancer tissue samples indicated that MB-COMT is expressed in breast tumors, with

expression that might exceed that of S-COMT, thus confirming a previous report that MB-COMT is expressed in breast cancer tissue (45) (Fig. 4B).

We next extended our studies to include two independent breast cancer case-control studies of similar ethnic composition. In the GENICA study, the protective effect of the two distal promoter SNPs was consistent with the Mayo results and that effect was observed for both the pre- and postmenopausal strata. However, the GESBC study, which consists mainly of premenopausal subjects, failed to replicate observations made for both Mayo Clinic and GENICA subjects. Heterogeneity of study results is a constant problem in association studies. However, our results showed that ORs differed significantly among the premenopausal women across all three studies. Differences in study design, composition of study populations, and underlying genetic and/or environmental factors might all contribute to these heterogeneous results. In addition, the possible effect of one or more undetected confounding factors in the GESBC study might have diluted the moderate protective role of these *COMT* variants. However, it should be emphasized that all of our functional observations were also compatible with a possible protective role for MB-COMT in estrogen-induced carcinogenesis. Subsequently, a complementary study was performed using the same Mayo Clinic study, and one of our two MB-COMT SNPs was also found to be associated with decreased mammographic density (46), further supporting our hypothesis since decreased mammographic density is associated with reduced breast cancer risk (47). These distal promoter SNPs should also be included in future neuropsychiatric studies of *COMT*.

In summary, we have identified two common, functionally significant polymorphisms located in the distal promoter of *COMT* which were associated with breast cancer risk reduction in 2 of 3 independent association studies – studies that included a total of 2327 cases and 2838 controls. Furthermore, laboratory-based functional genomic studies provided strong mechanistic support for these epidemiologic observations and suggested that the MB-form of the enzyme might contribute to a possible protective role for COMT in estrogen-induced breast carcinogenesis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Mrs. Luanne Wussow for her assistance with the preparation of this manuscript as well as Ursel Eilber and Belinda Kaspeseits (GESBC) for their technical assistance. The *COMT* gene resequencing data reported in this manuscript have been deposited in the National Institutes of Health-supported database PharmGKB with submission IDs: PS205182, PS205711 and PS206659.

Supported in part by NIH grants R01 GM28157, R01 GM35720 and U01 GM61388 (The Pharmacogenetics Research Network), a Breast Cancer Specialized Program of Research Excellence grant P50 CA166201 and a PhRMA Foundation Center of Excellence in Clinical Pharmacology Award. The GENICA study was supported by the Federal Ministry of Education and Research Germany grants 01KW9975/5, 01KW9976/8, 01KW9977/0, and 01KW0114, and the Robert Bosch Foundation of Medical Research, Stuttgart, Deutsches Krebsforschungszentrum, Heidelberg, Berufsgenossenschaftliches Forschungsinstitut für Arbeitsmedizin Bochum, and Medizinische Universitäts- und Poliklinik, Bonn, Germany. The GESBC study was supported by Deutsche Krebshilfe.

References

1. Smigal C, Jemal A, Ward E, et al. Trends in breast cancer by race and ethnicity: update 2006. *CA Cancer J Clin* 2006;56:168–83. [PubMed: 16737949]
2. Feigelson HS, Henderson BE. Estrogens and breast cancer. *Carcinogenesis* 1996;17:2279–84. [PubMed: 8968038]

3. Madigan MP, Ziegler RG, Benichou J, Byrne C, Hoover RN. Proportion of breast cancer cases in the United States explained by well-established risk factors. *J Natl Cancer Inst* 1995;87:1681–5. [PubMed: 7473816]
4. Henderson BE, Ross R, Bernstein L. Estrogens as a cause of human cancer: the Richard and Hinda Rosenthal Foundation award lecture. *Cancer Res* 1988;48:246–53. [PubMed: 2825969]
5. Cavalieri EL, Stack DE, Devanesan PD, et al. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. *Proc Natl Acad Sci U S A* 1997;94:10937–42. [PubMed: 9380738]
6. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. *N Engl J Med* 2006;354:270–82. [PubMed: 16421368]
7. Cavalieri E, Chakravarti D, Guttenplan J, et al. Catechol estrogen quinones as initiators of breast and other human cancers: implications for biomarkers of susceptibility and cancer prevention. *Biochim Biophys Acta* 2006;1766:63–78. [PubMed: 16675129]
8. Martucci C, Fishman J. Uterine estrogen receptor binding of catecholestrogens and of estetrol (1,3,5 (10)-estratriene-3,15alpha,16alpha,17beta-tetrol). *Steroids* 1976;27:325–33. [PubMed: 178074]
9. Van Aswegen CH, Purdy RH, Wittliff JL. Binding of 2-hydroxyestradiol and 4-hydroxyestradiol to estrogen receptors from human breast cancers. *J Steroid Biochem* 1989;32:485–92. [PubMed: 2542691]
10. Paul SM, Axelrod J. A rapid and sensitive radioenzymatic assay for catechol estrogens in tissues. *Life Sci* 1977;21:493–502. [PubMed: 904427]
11. Hoffman AR, Paul SM, Axelrod J. Catecholesterogen synthesis and metabolism by human breast tumors in vitro. *Cancer Res* 1979;39:4584–7. [PubMed: 498088]
12. Zhao Z, Kosinska W, Khmelnsky M, et al. Mutagenic activity of 4-hydroxyestradiol, but not 2-hydroxyestradiol, in BB rat2 embryonic cells, and the mutational spectrum of 4-hydroxyestradiol. *Chem Res Toxicol* 2006;19:475–9. [PubMed: 16544955]
13. Grossman MH, Emanuel BS, Budarf ML. Chromosomal mapping of the human catechol-O-methyltransferase gene to 22q11.1→q11.2. *Genomics* 1992;12:822–5. [PubMed: 1572656]
14. Tenhunen J, Salminen M, Lundstrom K, Kiviluoto T, Savolainen R, Ulmanen I. Genomic organization of the human catechol O-methyltransferase gene and its expression from two distinct promoters. *Eur J Biochem* 1994;223:1049–59. [PubMed: 8055944]
15. Lundstrom K, Tenhunen J, Tilgmann C, Karhunen T, Panula P, Ulmanen I. Cloning, expression and structure of catechol-O-methyltransferase. *Biochim Biophys Acta* 1995;1251:1–10. [PubMed: 7647086]
16. Jeffery DR, Roth JA. Characterization of membrane-bound and soluble catechol-O-methyltransferase from human frontal cortex. *J Neurochem* 1984;42:826–32. [PubMed: 6693904]
17. Weinshilboum RM, Raymond FA. Inheritance of low erythrocyte catechol-o-methyltransferase activity in man. *Am J Hum Genet* 1977;29:125–35. [PubMed: 848488]
18. Scanlon PD, Raymond FA, Weinshilboum RM. Catechol-O-methyltransferase: thermolabile enzyme in erythrocytes of subjects homozygous for allele for low activity. *Science* 1979;203:63–5. [PubMed: 758679]
19. Lachman HM, Papolos DF, Saito T, Yu YM, Szumlanski CL, Weinshilboum RM. Human catechol-O-methyltransferase pharmacogenetics: description of a functional polymorphism and its potential application to neuropsychiatric disorders. *Pharmacogenetics* 1996;6:243–50. [PubMed: 8807664]
20. Grossman MH, Szumlanski C, Littrell JB, Weinstein R, Weinshilboum RM. Electrophoretic analysis of low and high activity forms of catechol-O-methyltransferase in human erythrocytes. *Life Sci* 1992;50:473–80. [PubMed: 1542252]
21. Lotta T, Vidgren J, Tilgmann C, et al. Kinetics of human soluble and membrane-bound catechol O-methyltransferase: a revised mechanism and description of the thermolabile variant of the enzyme. *Biochemistry* 1995;34:4202–10. [PubMed: 7703232]
22. Shield AJ, Thomae BA, Eckloff BW, Wieben ED, Weinshilboum RM. Human catechol O-methyltransferase genetic variation: gene resequencing and functional characterization of variant allozymes. *Mol Psychiatry* 2004;9:151–60. [PubMed: 14966473]

23. Lavigne JA, Helzlsouer KJ, Huang HY, et al. An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer. *Cancer Res* 1997;57:5493–7. [PubMed: 9407957]
24. Millikan RC, Pittman GS, Tse CK, et al. Catechol-O-methyltransferase and breast cancer risk. *Carcinogenesis* 1998;19:1943–7. [PubMed: 9855007]
25. Thompson PA, Shields PG, Freudenheim JL, et al. Genetic polymorphisms in catechol-O-methyltransferase, menopausal status, and breast cancer risk. *Cancer Res* 1998;58:2107–10. [PubMed: 9605753]
26. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447:1087–93. [PubMed: 17529967]
27. Stram DO, Haiman CA, Hirschhorn JN, et al. Choosing haplotype-tagging SNPs based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered* 2003;55:27–36. [PubMed: 12890923]
28. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 2004;74:106–20. [PubMed: 14681826]
29. Shifman S, Bronstein M, Sternfeld M, et al. A highly significant association between a COMT haplotype and schizophrenia. *Am J Hum Genet* 2002;71:1296–302. [PubMed: 12402217]
30. Nackley AG, Shabalina SA, Tchivileva IE, et al. Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. *Science* 2006;314:1930–3. [PubMed: 17185601]
31. Olson JE, Ingle JN, Ma CX, et al. A comprehensive examination of CYP19 variation and risk of breast cancer using two haplotype-tagging approaches. *Breast Cancer Res Treat.* 2006
32. Justenhoven C, Hamann U, Pesch B, et al. ERCC2 genotypes and a corresponding haplotype are linked with breast cancer risk in a German population. *Cancer Epidemiol Biomarkers Prev* 2004;13:2059–64. [PubMed: 15598761]
33. Chang-Claude J, Eby N, Kiechle M, Bastert G, Becher H. Breastfeeding and breast cancer risk by age 50 among women in Germany. *Cancer Causes Control* 2000;11:687–95. [PubMed: 11065005]
34. Justenhoven C, Hamann U, Schubert F, et al. Breast cancer: a candidate gene approach across the estrogen metabolic pathway. *Breast Cancer Res Treat.* 2007
35. Slanger TE, Chang-Claude J, Wang-Gohrke S. Manganese superoxide dismutase Ala-9Val polymorphism, environmental modifiers, and risk of breast cancer in a German population. *Cancer Causes Control* 2006;17:1025–31. [PubMed: 16933053]
36. Wang L, Thomae B, Eckloff B, Wieben E, Weinshilboum R. Human histamine N-methyltransferase pharmacogenetics: gene resequencing, promoter characterization, and functional studies of a common 5'-flanking region single nucleotide polymorphism (SNP). *Biochem Pharmacol* 2002;64:699–710. [PubMed: 12167489]
37. Sladek-Chelgren S, Weinshilboum RM. Catechol-o-methyltransferase biochemical genetics: human lymphocyte enzyme. *Biochem Genet* 1981;19:1037–53. [PubMed: 7337686]
38. Raymond FA, Weinshilboum RM. Microassay of human erythrocyte catechol-O-methyltransferase: removal of inhibitory calcium ion with chelating resin. *Clin Chim Acta* 1975;58:185–94. [PubMed: 235384]
39. Friedenreich CM. Physical activity and breast cancer risk: the effect of menopausal status. *Exerc Sport Sci Rev* 2004;32:180–4. [PubMed: 15604938]
40. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;39:870–4. [PubMed: 17529973]
41. Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007;39:865–9. [PubMed: 17529974]
42. Mannisto PT, Kaakkola S. Catechol-O-methyltransferase (COMT): biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. *Pharmacol Rev* 1999;51:593–628. [PubMed: 10581325]

43. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BA. Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* 2002;31:33–6. [PubMed: 11984562]
44. Boudikova B, Szumlanski C, Maidak B, Weinshilboum R. Human liver catechol-O-methyltransferase pharmacogenetics. *Clin Pharmacol Ther* 1990;48:381–9. [PubMed: 2225698]
45. Tenhunen J, Heikkila P, Alanko A, Heinonen E, Akkila J, Ulmanen I. Soluble and membrane-bound catechol-O-methyltransferase in normal and malignant mammary gland. *Cancer Lett* 1999;144:75–84. [PubMed: 10503880]
46. Olson JE, Ji Y, Vachon CM, et al. Mammographic breast density and membrane-bound catechol-O-methyltransferase (COMT) genetic polymorphisms. *Proc Am Assoc Cancer Res*. 2008in press
47. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* 2006;15:1159–69. [PubMed: 16775176]

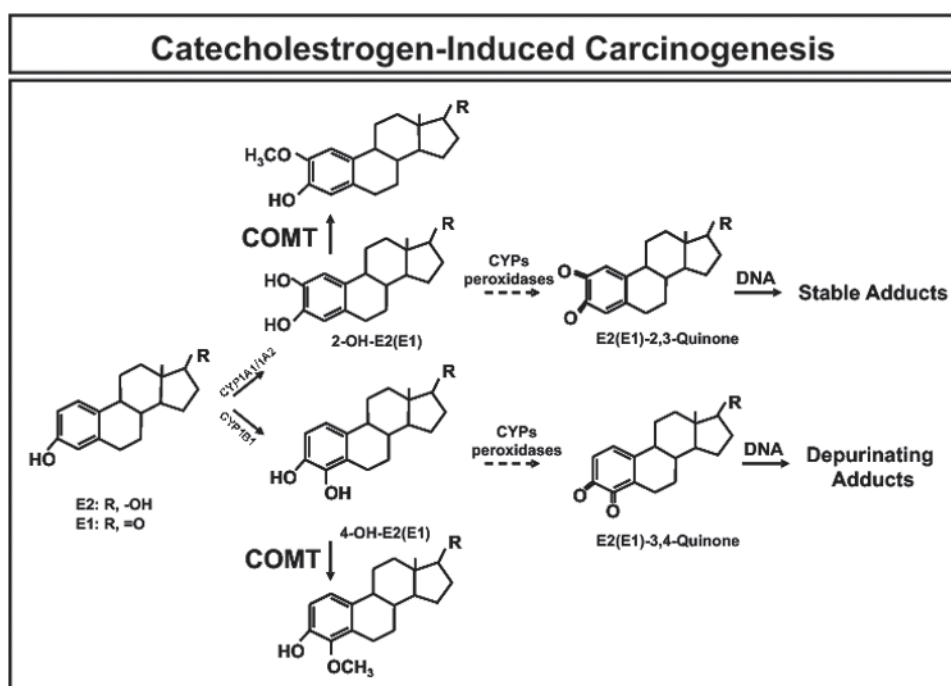


Figure 1.
Catecholesterogen-induced carcinogenesis.

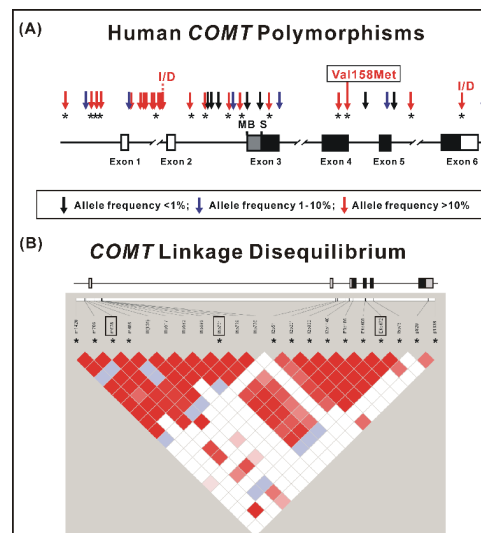


Figure 2.

Human *COMT* genetic polymorphisms and linkage disequilibrium. (A) *COMT* genetic polymorphisms in 60 Caucasian-American (CA) subjects. Arrows indicate the locations of polymorphisms. Black and gray rectangles represent coding exons, with the grey area specific for the MB-COMT ORF. Open rectangles represent non-coding exons. 'I/D' is insertion/deletion. Alterations in amino acid sequence resulting from nonsynonymous SNPs are boxed. Polymorphisms labeled with an asterisk were genotyped in the Mayo Clinic study. (B) *COMT* linkage disequilibrium displayed by the use of Haploview 3.3. Polymorphisms with MAF <5% were excluded from the analysis. The coloring scheme is the standard Haploview (D'/LOD).

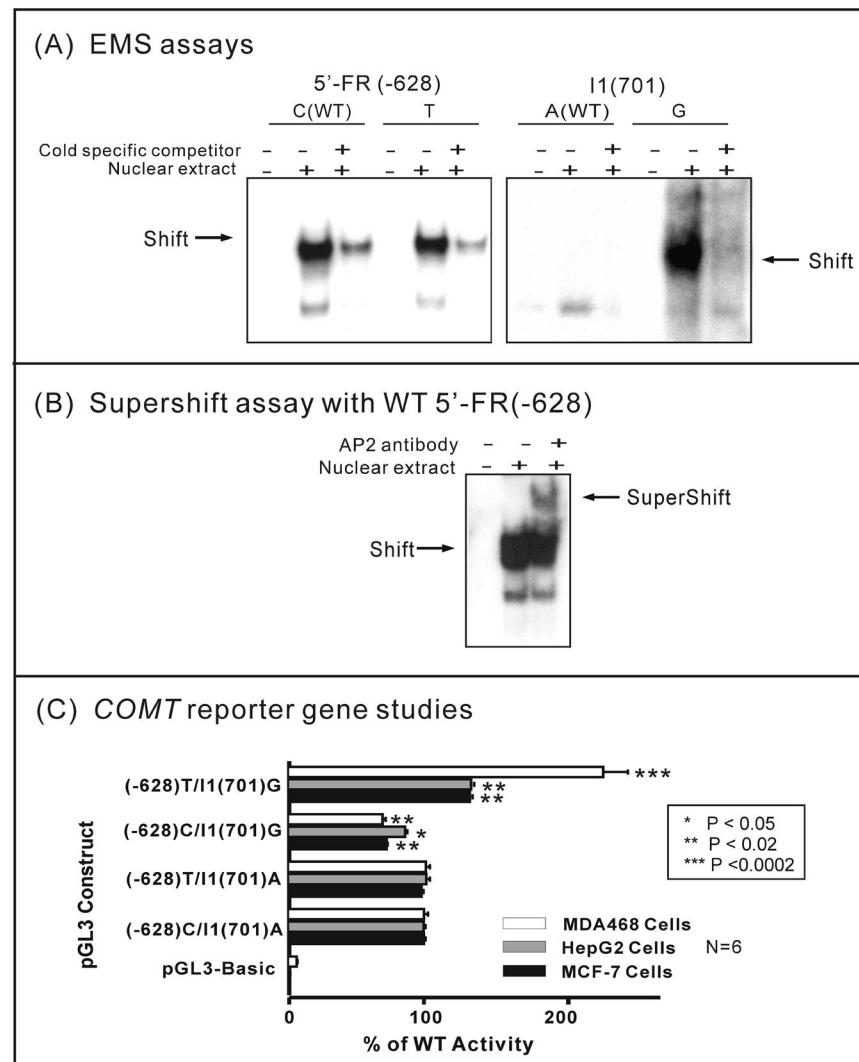


Figure 3. COMT functional genomic studies. (A) EMS assays for 5'-FR(-628) and I1(701) SNPs. MCF-7 cell nuclear extract was incubated with 32 P-labelled oligonucleotides. (B) Supershift assay for WT 5'-FR(-628). A rabbit polyclonal antibody against the human AP-2 α was incubated with 32 P-labelled WT 5'-FR (-628) oligonucleotides and MCF-7 cell nuclear extract. (C) Reporter gene studies. Activities of luciferase reporter gene constructs containing different combinations of the two loci, 5'-FR(-628) and I1(701), are shown as a percentage of the level of relative luciferase activity for the WT construct, (-628)C/I1(701)A. Each bar represents the mean \pm SEM of 6 independent transfections.

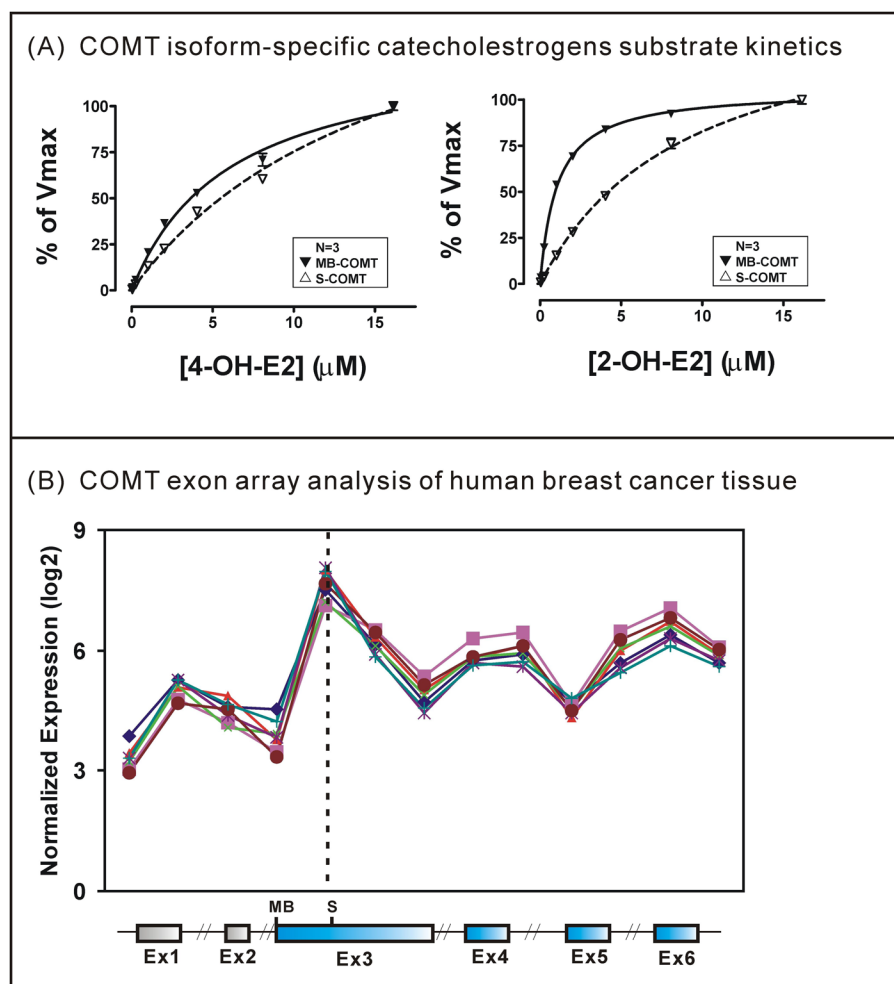


Figure 4.

COMT isoform-specific substrate kinetics and expression in breast cancer tissue. (A) Substrate curves for the methylation of 4-OH-E2 and 2-OH-E2 catalyzed by either MB-COMT or S-COMT. Each point is the mean \pm SEM for 3 determinations. (B) Exon array analysis of COMT in human breast cancer tissue. Each point represents the normalized expression value for that probe set. The *COMT* gene structure is also shown, with blue boxes representing ORF and grey boxes non-coding exons.

Association of SNPs for Mayo Clinic breast cancer cases and controls, stratified on menopausal status. The three SNPs that were genotyped in the two replication studies are “boxed”.

Table 1

Menopausal Status	SNP	rs#	MAF		Adjusted - Matching (1)			Adjusted - Covariates (2)			
			Cases	Controls	OR (allelic)	95% CI	p-value	Permutation p-value	OR (allelic)	95% CI	p-value
Premenopausal			N = 264								
			N = 193								
	5'FR(-1420)	rs2075507	0.46	0.43	1.07	0.81 - 1.41	0.619	0.999	1.07	0.81 - 1.42	0.635
	5'FR(-706)	rs4646310	0.19	0.18	1.03	0.72 - 1.47	0.869	1.000	1.05	0.73 - 1.50	0.791
	5'FR(-628)	rs2020917	0.22	0.30	0.70	0.52 - 0.95	0.022	0.201	0.70	0.52 - 0.96	0.024
	5'FR(-485)	rs13306278	0.11	0.14	0.79	0.53 - 1.17	0.241	0.966	0.77	0.52 - 1.16	0.217
	Intron1(701)	rs737865	0.23	0.32	0.68	0.51 - 0.92	0.011	0.102	0.69	0.51 - 0.92	0.013
	Intron2(51)	rs165656	0.44	0.48	0.86	0.66 - 1.13	0.274	0.975	0.88	0.67 - 1.16	0.379
	Intron2(201)	rs165722	0.44	0.48	0.87	0.67 - 1.14	0.316	0.985	0.90	0.69 - 1.18	0.453
	Intron2(832)	rs3810595	0.35	0.41	0.79	0.60 - 1.03	0.080	0.662	0.81	0.61 - 1.06	0.130
	Intron2(1140)	rs6269	0.36	0.42	0.78	0.60 - 1.02	0.072	0.666	0.82	0.62 - 1.08	0.153
	Exon3(186)	rs4633	0.44	0.48	0.88	0.68 - 1.15	0.358	0.991	0.91	0.70 - 1.20	0.509
	Exon4(408)	rs4818	0.34	0.40	0.78	0.59 - 1.03	0.081	0.703	0.81	0.61 - 1.07	0.134
	Exon4(472)	rs4680	0.42	0.46	0.89	0.68 - 1.17	0.411	0.992	0.91	0.69 - 1.20	0.499
	Intron5(75)	rs4646315	0.23	0.19	1.18	0.84 - 1.67	0.336	0.795	1.14	0.81 - 1.62	0.454
	3'UTR(820 indel)	rs362204	0.24	0.25	0.97	0.69 - 1.37	0.882	1.000	0.99	0.70 - 1.41	0.976
	3'FR(1338)	rs165599	0.29	0.30	0.98	0.73 - 1.32	0.884	1.000	0.99	0.73 - 1.35	0.972
Postmenopausal			N = 454								
			N = 521								
	5'FR(-1420)		0.43	0.43	1.01	0.83 - 1.23	0.917	1.000	0.98	0.81 - 1.20	0.853
	5'FR(-706)		0.18	0.17	1.05	0.82 - 1.35	0.714	1.000	1.07	0.83 - 1.39	0.594
	5'FR(-628)		0.26	0.27	0.94	0.76 - 1.17	0.600	1.000	0.97	0.78 - 1.20	0.765
	5'FR(-485)		0.14	0.15	0.89	0.67 - 1.17	0.389	1.000	0.90	0.68 - 1.18	0.436
	Intron1(701)		0.27	0.27	0.97	0.78 - 1.21	0.795	1.000	1.00	0.80 - 1.25	0.991
	Intron2(51)		0.47	0.48	0.97	0.81 - 1.17	0.765	1.000	0.98	0.81 - 1.19	0.863
	Intron2(201)		0.47	0.48	0.94	0.79 - 1.13	0.524	1.000	0.96	0.79 - 1.15	0.643
	Intron2(832)		0.37	0.40	0.89	0.74 - 1.07	0.210	0.962	0.90	0.74 - 1.09	0.279
	Intron2(1140)		0.37	0.41	0.84	0.70 - 1.02	0.073	0.657	0.86	0.71 - 1.04	0.109
	Exon3(186)		0.47	0.49	0.92	0.77 - 1.11	0.383	1.000	0.93	0.77 - 1.12	0.444
	Exon4(408)		0.36	0.39	0.90	0.75 - 1.09	0.296	1.000	0.91	0.75 - 1.11	0.364
	Exon4(472)		0.46	0.47	0.95	0.78 - 1.14	0.567	1.000	0.95	0.78 - 1.15	0.608
	Intron5(75)		0.20	0.18	1.15	0.90 - 1.46	0.263	0.975	1.14	0.89 - 1.46	0.300
	3'UTR(820 indel)		0.26	0.27	0.95	0.76 - 1.18	0.619	1.000	0.95	0.76 - 1.18	0.626
	3'FR(1338)		0.32	0.30	1.03	0.84 - 1.26	0.806	1.000	1.04	0.84 - 1.28	0.742

(1) = Adjusted for matching variables of age and area.

(2) = Adjusted for age, area, parity/age at first child birth, and physical activity.

Table 2
Association of SNPs in the two replication studies, stratified on menopausal status.

Group	Menopausal Status	SNP	MAF		Unadjusted			Adjusted - Covariates		
			Cases	Controls	OR (allelic)	95% CI	p-value	OR (allelic)	95% CI	p-value
GENICA	Premenopausal	5'FR(-628) Intron1 (701)	N = 249 0.26 0.24	N = 234 0.30 0.28	0.79 0.79	0.59 – 1.05 0.59 – 1.06	0.106 0.114	see footnote ⁽¹⁾ 0.78 0.78	0.59 – 1.05 0.58 – 1.05	0.104 0.106
		Exon4 (472)	0.47	0.48	0.98	0.76 – 1.26	0.884	0.98	0.76 – 1.26	0.867
	Postmenopausal	5'FR(-628) Intron1 (701)	N = 758 0.25 0.23	N = 763 0.28 0.26	0.85 0.84	0.72 – 1.00 0.72 – 0.99	0.044 0.042	see footnote ⁽¹⁾ 0.85 0.85	0.72 – 1.00 0.72 – 1.01	0.053 0.061
		Exon4 (472)	0.49	0.47	1.07	0.93 – 1.23	0.351	1.07	0.93 – 1.23	0.367
GESBC	Premenopausal	5'FR(-628) Intron1 (701)	N = 438 0.30 0.30	N = 877 0.28 0.29	1.08 1.04	0.90 – 1.28 0.87 – 1.24	0.412 0.655	see footnote ⁽²⁾ 1.09 1.05	0.91 – 1.31 0.88 – 1.26	0.337 0.587
		Exon4 (472)	0.50	0.49	1.03	0.88 – 1.21	0.719	1.03	0.87 – 1.22	0.740
	Postmenopausal	5'FR(-628) Intron1 (701)	N = 124 0.30 0.30	N = 208 0.29 0.29	1.05 1.04	0.74 – 1.49 0.73 – 1.47	0.783 0.835	see footnote ⁽²⁾ 1.06 1.04	0.74 – 1.51 0.73 – 1.48	0.743 0.827
		Exon4 (472)	0.46	0.49	0.87	0.63 – 1.20	0.399	0.87	0.63 – 1.21	0.400

⁽¹⁾ = Adjusted for HRT use and physical activity.

⁽²⁾ = Adjusted for physical activity and parity/age at first child birth.